

Variation in exploration and exploitation in group decision-making: Evidence from immersive simulations of major incident emergencies

Byron Wilkinson | Sabrina R. Cohen-Hatton | Robert C. Honey

School of Psychology, Cardiff University, Cardiff, UK

Correspondence

Robert C. Honey, School of Psychology, Cardiff University, Cardiff CF10 3AT, UK.
Email: honey@cardiff.ac.uk

Funding information

Economic and Social Research Council, Grant/Award Number: ES/M500422/1

Abstract

Multi-agency groups are brought together to make strategic decisions in response to major incident emergencies. Here, we investigated decision-making processes in 18 multi-agency groups who were video recorded while engaged in simulated major incident emergencies involving a potential need to evacuate individuals from the location of the incident. Three general categories of decision-making activity were used to code the videos: situation assessment (SA), plan formulation (PF) and plan execution (PE). Analysis of the transitions between these decision-making activities showed that there were marked between-group departures from normative models of decision-making, which predict an orderly transition from SA to PF and then from PF to PE. These departures appeared to reflect between-group differences in the tendency to explore information (evident in reciprocal transitions between SA and PF) or exploit information (apparent in transitions to and from SA and PF to PE). Moreover, the tendency to explore but not exploit information was associated with the number of transitions to critical decisions (i.e. to evacuate individuals from the location of the incident).

KEYWORDS

decision making, planning, scenario

1 | GROUP DECISION-MAKING AT MAJOR INCIDENTS

Many important decisions are made by groups of individuals on behalf of society. For example, when there is a major incident or disaster in the UK, a group of senior representatives assemble from the local emergency services, civil resource organizations, health agencies and government. These Strategic Coordinating Groups (SCGs) are charged with making decisions that help to minimize the immediate and ongoing impacts of major incidents including societal and economic disruption. Recent major incidents in the UK

include terrorist attacks (e.g. Manchester Arena, Borough Market and Houses of Parliament), the Grenfell Tower fire and natural disasters including severe weather events. The scale of such incidents requires SCGs to consider the incident in its wider context, define and communicate an overarching policy and strategy and monitor progress towards defined objectives. This role extends beyond the initial response to the incident and includes formulating a media and communication strategy as well as horizon scanning to facilitate the recovery stage of an incident. Notwithstanding, the heroism shown by the frontline responders attending major incidents (e.g. firefighters, paramedics and police officers), the failure of different agencies

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Journal of Contingencies and Crisis Management* published by John Wiley & Sons Ltd.

to work together effectively at a strategic level is an ongoing issue in the UK and elsewhere (e.g. Alison & Crego, 2008; Comfort, 2007; Flin, 1996; House et al., 2014; Pollock, 2013).

2 | NORMATIVE MODELS FOR GROUP DECISION-MAKING AT MAJOR INCIDENTS

Decision-making in SCGs and the UK emergency services more broadly is guided by the Joint Decision Model (JDM; see Cabinet Office, 2013; JESIP Doctrine). This model, based on the UK Police National Decision Model (NDM; <https://www.app.college.police.uk/app-content/national-decision-model/the-national-decision-model/#the-model>), includes a prescribed sequence of five categories of decision-making activity: gather information and intelligence; assess risks and develop a working strategy; consider powers, policies and procedures; identify options and contingencies and take action and review what happened next, at which point the guidance returns to gather information and intelligence. A recent analysis of group decision-making at immersive simulated major incidents revealed marked departures from the use of the JDM, but also significant between-group variation in the transitions between the five categories of decision-making activities (Wilkinson et al., 2019; see also, Waring et al., 2020). Thus, there was limited consideration of options and contingencies across all groups, and there was marked between-group variation in the transitions between the activities that they engaged in: Some groups tended to move back and forth between, on the one hand, gathering information and intelligence, and on the other, assessing risks and developing a working strategy. In other groups, the reciprocal sequences involved taking action, and both gathering information and intelligence, and assessing risks and developing a working strategy. These results indicate that SCGs do not follow the guidance enshrined in the JDM. However, characterizing the decision-making processes in SCGs through the lens and bespoke categories of this model makes it difficult to understand the process of decision-making in the SCGs observed by Wilkinson et al. (2019). For example, the JDM like the NDM separates components of decision-making that one might consider, on an a priori basis, to be intimately related (e.g. assess risks and develop a working strategy, and identify options and contingencies); similarly, the model combines components of decision-making that one might consider, on an a priori basis, to be separate (e.g. take action and review what happened next).

A more general analysis of decision-making processes, which has been used in other operational contexts, has used three categories: situation assessment (SA; e.g. "We need more information about how injuries were sustained."), plan formulation (PF; "How will we transport and shelter affected people?") and plan execution (PE; e.g. "Initiate mutual aid plan."). This coding system was developed by van den Heuvel, Alison and Power (2012; see also, van den Heuvel et al., 2014). Here too, one could assume that decision-making follows a sequence (from SA to PF to PE and back to SA), which echoes normative models of individual decision-making across a variety of

domains (e.g. Dewey, 1933; see also, Groenendaal & Helsloot, 2016; Lipshitz & Bar-Ilan, 1996; van den Heuvel et al., 2012, van den Heuvel et al., 2014). This coding system has been used to reveal decision-making processes in firefighters, for whom information gathering (i.e. situation assessment, SA) is often followed by courses of action (i.e. plan execution, PE) without any apparent mediation by a stage of evaluation or plan formulation (i.e. PF; see Cohen-Hatton et al., 2015; Cohen-Hatton & Honey, 2015; see also, Klein, 1998; Klein, Calderwood & MacGregor, 1989). Here, we applied this more general form of analysis to the video footage of 18 multi-agency groups originally analysed by Wilkinson et al. (2019; see also, Waring et al., 2019) in order to assess a recent theoretical analysis of group decision-making, which is consistent with aspects of the naturalistic decision-making approach.

3 | EXPLORATION AND EXPLOITATION IN GROUP DECISION-MAKING

Bang and Frith (2017) presented a synthesis of the evidence concerning the benefits, as well as the pitfalls, of making decisions in groups rather than individually. Their overarching (Bayesian) theoretical analysis, which involved how the past experience of group members is integrated with new information to affect group decisions, is broadly consistent with the naturalistic decision-making approach. A central component of this approach, developed by Klein (1993, 2003, 2008; see also, Klein et al., 2003; Hutchins, 1995a, 1995b), was based on just this kind of interaction: how previous experience primes decisions in the face of uncertain information (i.e. recognition-primed or intuitive decision-making; see also Doya, 2008; Gigerenzer, 2007; Gureckis & Goldstone, 2006; Salas et al., 2010; Tversky & Kahneman, 1974). In the case of multi-agency groups responding to major incidents, a natural assumption is that bringing together representatives from the relevant agencies might increase their ability to work together effectively (for a discussion of distributed vs. co-localized multi-agency command structures, see Power & Alison, 2017). However, Bang and Frith (2017) argued that group decision-making might also benefit from the combination of different types of individual decision maker: specifically the combination of *explorers* and *exploiters*.

Explorers can be characterized as sampling the available information and decision space in order to select the optimal decision, whereas exploiters commit to a course of action without such a thoroughgoing analysis and based on the prior success of that action (see Frank et al., 2009; see also Badre, Doll, Long & Frank, 2012; Cohen, McClure & Yu, 2007; Daw, O'Doherty, Dayan, Seymour & Dolan, 2006). For example, in the "observe or bet" task the decision maker has two options that yield rewards or losses according to a predetermined bias (Tversky & Edwards, 1966). They can choose to "observe" the outcome of a trial and gain information but accrue no rewards or losses or "bet" and accumulate rewards or losses that are only revealed at the end of the task. The observe trials represent an assay of exploration and the bet trials an assay of exploitation, and

their use varies across individuals. Exploration can be linked with one way in which the term *decision inertia* is used: Where individuals and groups defer taking action and instead continue to seek additional (redundant) information (Alison et al., 2015). While exploitation can be linked to another way in which the same term is used: the tendency for individuals to repeat past choices irrespective of the current evidence (Akaishi et al., 2014).

There are clearly pitfalls associated with being either an explorer (who might not reach a decision in a timely fashion) or an exploiter (who might quickly reach the wrong decision). Bang and Frith (2017; p. 7) state: "A mixture of such diverse individuals can create advantages for the group." This claim has clear practical implications for assembling effective groups in a variety of contexts, including major incidents (cf. Polzer et al., 2002; Roberge & van Dick, 2010). However, to the best of our knowledge, the tendency for groups to engage in exploration and exploitation during decision-making has not been investigated. Certainly, the selection of individuals that come together to respond to major incidents is not based on any formal assessment of their individual approaches to decision-making. Consequently, if the decision-making style of group members contributes to the tendency of groups to engage in exploration or exploitation, then one would predict variation in these processes across different groups.

4 | ASSESSING GROUP DECISION-MAKING PROCESSES IN IMMERSIVE SIMULATIONS

Immersive simulation learning environments enable the components of major incidents to be simulated and provide a basis for training exercises and research on group decision-making (see Alison et al., 2013). These Hydra environments (Crego, 1996) consist of a "syndicate room" for group members that is equipped with a large screen projector, PC and CCTV. The PC runs a communication interface that is permanently displayed on the projector screen and delivers information updates ("injects") and tasks to the groups. In this way, table-top exercises are lifted to a new level of realism, and the environments provide a platform for investigating how the provision of different forms of information affect group behaviour. They also enable the processes of group decision-making to be investigated in situ, through a real-time analysis of the recordings of real multi-agency groups faced with reproducible simulated major incidents. In fact, Hydra is used by the emergency services across the UK (<https://www.hydrafoundation.org>). This approach has the potential to separate variation in group decision-making that might reflect the idiosyncrasies of different real incidents from those based on the groups that deal with them. It permits important theoretical issues concerning group decision-making to be systematically evaluated.

Here, the sequences of decision-making activities (situation assessment, SA; plan formulation, PF; and plan execution, PE) derived from the recordings allowed us to assess the extent to which the normative model is followed in these important groups (i.e. SA-PF-PE; cf. Burke et al., 2006). We also used these sequences to assess whether

any departure from a normative standard reflect between-group differences in exploration and exploitation (cf. Bang & Frith, 2017). This was achieved through further analysis of the differences in the sequences of decision-making activities across groups: Exploration should be evident as repeatedly moving between situation assessment and plan formulation, whereas exploitation should be evident in transitions between situation assessment and plan execution, and between plan formulation and plan execution. Any such differences in exploration and exploitation across groups should be reflected, other things being equal, in the number of transitions before a critical decision is made: Groups who explore should take more steps to reach a critical decision than those who exploit.

5 | THE SCENARIOS AND APPROACH

We recorded decision-making activities of real multi-agency groups faced with two simulated major incidents: a large-scale chemical fire at an industrial site (Scenario 1) and a crash between a passenger train and a truck carrying a hazardous substance (Scenario 2). The scenarios were dynamic and unfolding in time, and the groups were located in Hydra immersive simulation suites that allowed information to and from the groups to be controlled and recorded (Alison et al., 2013; Crego, 1996). These suites are modelled on special operations rooms in which such groups meet in the context of real incidents. The groups consisted of senior representatives from the relevant agencies (e.g. local emergency services, civil resource organizations, health boards and government). In both scenarios, the critical decision was whether or not to evacuate local residents. However, in Scenario 1 the groups were tasked with developing a communications strategy and the decision to evacuate was an implicit component of the task, whereas in Scenario 2 the groups were explicitly tasked with deciding whether to evacuate local residents. The recordings of the meetings were coded as a continuous sequence of three decision-making activities: SA, PF and PE (examples from Scenarios 1 and 2 are presented in Table 1). As in Wilkinson et al. (2019), the coding was conducted independently of the group members who contributed to the decision-making activity. We then assessed the patterns of transitions between the successive categories. The first question was whether groups followed the normative model and consistently moved from situation assessment

TABLE 1 Examples of situation assessment (SA), plan formulation (PF) and plan execution (PE)

Scenario 1

SA "What is the information on the state of the fire and the risk of explosion?"

PF "We need to develop a media strategy."

PE "We will not invoke the evacuation plan at this time."

Scenario 2

SA "Do we know how many people might need to be evacuated?"

PF "A transport plan is needed."

PE "Invoke the Mass Fatalities plan."

to plan formulation to plan execution and then further situation assessment in making decisions across the course of the meetings. We then assessed whether deviations from the normative model between groups reflected differences in the tendency to explore or exploit information. As already noted, it was anticipated that while exploration would be evident in transitions between situation assessment and plan formulation (i.e. SA-PF and PF-SA), exploitation would be reflected in transitions to and from plan execution (most obviously between SA-PE and PF-PE). Finally, we examined whether such differences were related to the number of transitions that each group took to arrive at the critical decision (i.e. whether or not to evacuate local residents).

6 | METHOD

6.1 | Participants

Eighteen multi-agency groups attended 2-day national training and exercise events in Scenario 1 (nine groups; mean group size = 8 [range: 6–10]; female: male = 1:1.84) and Scenario 2 (nine groups; mean group size = 8 [range: 7–10]; female: male = 1:2.61). The groups included senior members of the blue-light emergency services (Ambulance, Fire and Rescue, and Police), but included differing numbers of other responding services. Scenario 1 involved training events in Wales from 2015 (called Exercise Wales Gold 1), while those in Scenario 2 involved training events in Wales from 2016 (called Exercise Wales Gold 2). Each training event consisted of opportunity samples of participants who had applied and were selected by their respective agencies to take part on the basis of prior or future involvement in multi-agency groups at major incident emergencies. Participants were employed by the emergency services (Ambulance Service, Fire and Rescue Service, and Police), Health Boards, Local and Regional Government, Natural Resources Wales and Public Health Wales, and ranged in seniority from chief executive level to a recently promoted Police Chief Superintendent. These groups consisted of participants who could, if an immediate need arose, be part of multi-agency groups faced with a major incident. Their experience varied considerably: The majority had either served as a member of such groups or had undertaken some relevant training to prepare them for such a role, while others had received no formal training for the role. The participants provided informed consent for their participation in accordance with ethical approval through Cardiff University.

6.2 | Procedure

At the start of Scenarios 1 and 2, all participants were given an interactive training session that lasted for approximately one hour. The information covered in this meeting included the consistent and nationally recognized UK emergency command control and coordination structures (Cabinet Office, 2013), the decision-making role of the group and the use of the nationally endorsed model for making collective decisions in an emergency (i.e. the Joint Decision Model).

Across the two days of the training event, the groups took part in timed meetings that were approximately 45–60 min, during which they made decisions in response to the evolving incident. The virtual timeline of the scenarios extended from the afternoon of Day 1, when the incident was declared, to 3 months later through three intermediate time points (evening of Day 1, Day 2 and 1 week later). The scenarios were delivered using Hydra immersive simulation systems (Crego, 1996). As already noted, Hydra provides a “syndicate room” for each group that is equipped with a large screen projector, PC, wireless keyboard and mouse, printer and CCTV. The PC ran a communication interface that was permanently displayed on the projector screen and delivered information updates (“injects”) and tasks to the groups (which were delivered at approximately the same time to all groups). Both scenarios were multi-faceted, dynamic and involved time pressure. The scenarios were developed, so that there were no explicitly correct or incorrect critical decisions and that all agencies would be engaged. The scenarios were managed by control room staff, who delivered scripted updates for each of the two scenarios at pre-defined times through the introduction of tasks, video and audio clips and printed documents. Exercise control staff also received and responded to all the written communications sent by the groups using the Hydra communicator.

Scenario 1 started with an update from the Police tactical commander through a video link. This update stated that there was a large-scale chemical fire at an industrial site, and that the nearby road network and railway line had been closed due to the resulting plume of smoke. The scenario then developed into a serious environmental and economic incident, with media impacts that required decisions on longer-term recovery issues involving health impacts, housing, decontamination and economic recovery. Scenario 2 also started with a video update from the Police tactical commander. However, in this case the scenario involved a crash between a passenger train and a truck, which was carrying a hazardous substance. The crash caused many fatalities and injuries to passengers on the train. Within the first hour of the incident a fire ignited, burning the hazardous substance and sending a toxic plume of smoke over a residential area. The analysis for Scenarios 1 and 2 was conducted on the critical second group meeting. In this meeting, there was time pressure and the groups were required to make critical decisions: In Scenario 1, the groups were tasked with providing direction to those involved in tactical operations (including whether it would be necessary to evacuate local residents) and what their media strategy would be; and in Scenario 2, they were explicitly tasked with deciding whether or not to evacuate the nearby caravan site, under conditions where the resources were not available to evacuate everyone and the toxic effects of the plume were unclear.

6.3 | Coding of activity

The audio–video recordings, from either the Hydra CCTV system or a GoPro camera placed on each group table, were coded using the categories: situation assessment, plan formulation and plan

execution. These categories were coded at the level of the group (i.e. independently of the individual contributor) and noted on a spreadsheet for later analysis. Isolated irrelevant comments and those that were not parts of the group discussions (e.g. informal asides, which were infrequent) were excluded from the analysis. The coding was conducted on two separate occasions (by BW), which resulted in a small number of the activities (<5%) being re-coded. An independent assessor (RCH) then confirmed the reliability of the coding on a sample of 30 observations from each study ($\approx 95\%$ agreement with BW). A lag sequential analysis (Sackett, 1979; see also, O'Connor, 1999) was used to derive the primary data of interest: the sequences of transitions between the decision-making activities in the group meetings. In such an analysis, repetitions of the same category are removed. The lag sequential analysis stopped at the end of the group meetings.

6.4 | Analytic approach

To assess whether there were differences between the frequencies of the different categories of decision-making activities, we used within-subjects ANOVAs with post-hoc t tests using a Bonferroni correction for multiple comparisons. As already noted, a lag sequential analysis was used to characterize the six two-step transitions between these categories (i.e. SA–PF, PF–SA, PF–PE, PE–PF, PE–SA and SA–PE). A principal-components analysis (PCA) was conducted to investigate whether or not there was structure within the patterns of variation across the six transitions: whether variation in some of the sequences was related to variation in the other sequences. The relationships between components from the PCA (indicative of exploration and exploitation) and the number of transitions to the critical decision (to evacuate or not) was assessed using bivariate correlations. A complementary analysis was conducted in which the multi-agency groups were separated into two groups, based on their exploration-exploitation bias, and then a t test was conducted on the number of transitions to the critical decision. Supplementary analyses were conducted in order to understand the relationships between exploration, exploitation, the overall number of transitions during the sessions and multi-agency group size, and to assess alternative analyses for the relationships between exploration, exploitation and the critical decision.

7 | RESULTS

7.1 | Overall analysis of the transitions

Each pair of bars in the upper panel (from Scenario 1) and lower panel (from Scenario 2) of Figure 1 shows the mean number of transitions from a given initiating category (e.g. situation assessment, SA) to the succeeding categories (plan formulation, PF, or plan execution, PE). If groups were following the normative sequential model for each decision, then the left bar from each pair should be higher

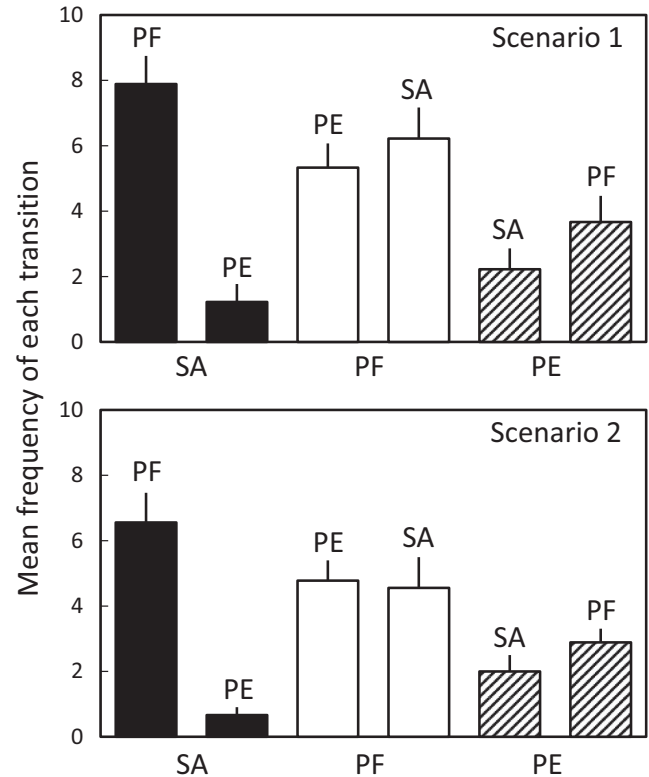


FIGURE 1 The mean frequencies (+SEM) of transitions from each of the categories (situation assessment, SA; plan formulation, PF and plan execution, PE) to the other two categories. The category labels below the panels (e.g. SA) indicate the first element of the sequences involving the categories denoted by equivalently coloured bars immediately above them (i.e. PF and PE). The results from Scenarios 1 and 2 are depicted in the upper and lower panels, respectively

than the right: Situation assessment (SA) should be followed by plan formulation (PF), plan formulation (PF) by plan execution (PE) and plan execution (PE) by situation assessment (SA). The impression gained from examining Figure 1 is that the pattern of transitions was similar in Scenarios 1 and 2, and in both cases was not that predicted on the basis of normative models: While situation assessment (SA) was more likely to be followed by plan formulation (PF) than plan execution (PE), plan formulation (PF) was equally likely to be followed by plan execution (PE) and situation assessment (SA), and plan execution (PE) was less likely to be followed by situation assessment (SA) than (PF).

ANOVA with scenario (1 or 2) as a between-subjects factor and initiating category (SA, PF and PE) and succeeding category (normative or other) as within-subjects factors, confirmed these impressions. There was no effect of scenario, $F(1, 16) = 2.39, p > .14, \eta_p^2 = .13$, but there were main effects of initiator category, $F(2, 16) = 29.87, p < .001, \eta_p^2 = .65$ and succeeding category, $F(1, 16) = 22.86, p < .001, \eta_p^2 = .59$. There was also an interaction between the initiator and succeeding category, $F(2, 32) = 23.80, p < .001, \eta_p^2 = .60$. There were no other interactions ($F_s < 1$). Pairwise comparisons confirmed that transitions from SA were more likely to be to PF than PE ($t[17] = 8.19, p < .001, d = 3.13$), but transitions from PF

were no more likely to be to PE than SA ($t[17] = -0.38, p = .709, d = -0.08$), and transitions from PE were more likely to be to PF than SA ($t[17] = -2.12, p < .05, d = -0.47$).

7.2 | Group differences in exploration and exploitation

The overall analysis of the results presented in Figure 1 could give the impression that while the pattern of results did not conform to what might be predicted on the basis of a normative model of decision-making (i.e. SA-PF, PF-PE), the groups were adopting a relatively consistent approach to decision-making. However, this impression would not be accurate. A PCA was conducted on the frequencies of the 6 possible transitions between these three categories (i.e. SA-PF, PF-SA, PF-PE, PE-SA and SA-PE) to investigate whether or not variation in some of the sequences was related to variation in the other sequences. This analysis converged in three iterations and used a varimax rotation and Kaiser Normalization. It revealed two transition classes with eigenvalues of >1 : Class 1 can be labelled “exploration” and involved the transitions: SA-PF, PF-SA (both loadings $>.93$); and Class 2 can be labelled “exploitation,” with all transitions involving PE: SA-PE, PE-SA, PF-PE and PE-PF (all loadings $>.70$). There were no cross loadings $>\pm.28$. These two classes accounted for 76% of the variance in the frequencies of the six transitions between the three categories.

7.3 | Relationship between exploration, exploitation and critical decisions

Figure 2 shows the relationships between the PCA loadings for exploration (upper panel) and exploitation (lower panel), and the numbers of transitions to the critical decision to evacuate (open symbols) or not (closed symbols) irrespective of the nature of those transitions: The higher the loading, the more a given group can be said to be engaging in exploration (SA-PF, PF-SA; upper panel) or exploitation (SA-PE, PE-SA, PF-PE, PE-PF; lower panel). It is important to note that while the exploration and exploitation loadings are based on the nature of the transitions between SA, PF and PE, the number of transitions to the critical decision is coded independently of the nature of the transitions. There were 13 groups that reached the critical decision to evacuate or not within the session. Inspection of the upper panel shows that there was a correlation between exploration and the number of transitions to the critical decision, $r(13) = .55, p < .05$, with higher loadings associated with more transitions to the decision. However, there was no relationship between exploitation and the number of transitions to the critical decision, $r(13) = -.22, p > .46$. Figure 3 depicts the results from a complementary analysis, where the 13 groups were divided based on their decision-making bias (exploration loading minus exploitation loading) to form group Explore ($n = 6$; mean bias = 1.22; $SEM = 0.38$) and group Exploit ($n = 7$; mean bias = -0.92 ; $SEM = 0.37$). There was a significant

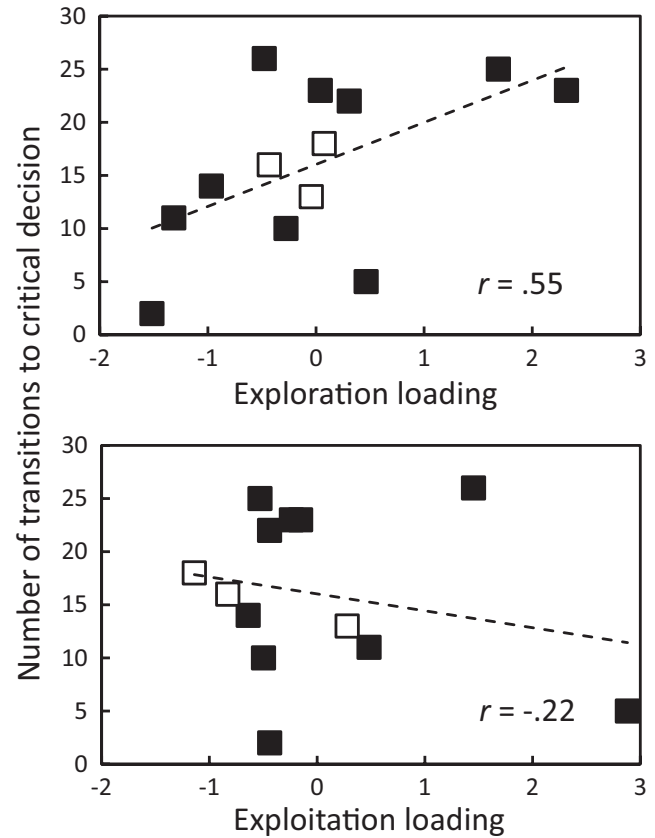


FIGURE 2 The relationship between exploration and the number of transitions to arrive at the critical decision (upper panel) and between exploitation and the number of transitions to the critical decision (lower panel). The total number of groups who came to a decision was 13, with 10 deciding to evacuate (filled symbols) and 3 deciding not to evacuate (open symbols)

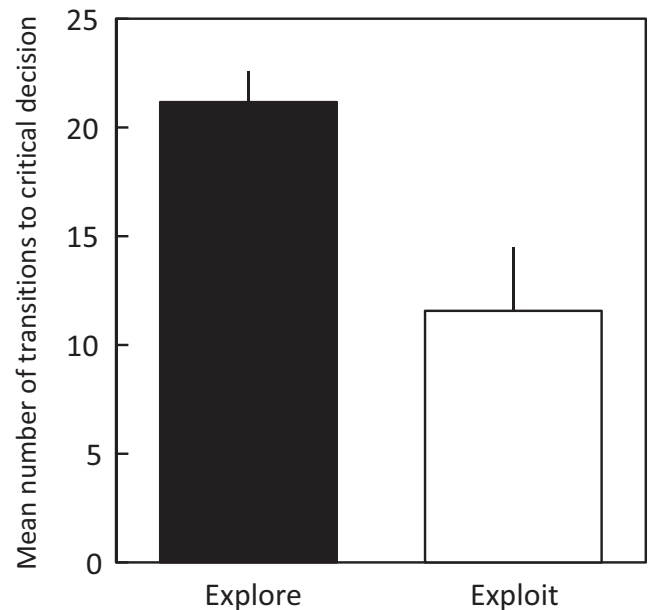


FIGURE 3 The mean number of transitions ($\pm SEM$) to reach the critical decision in groups Explore and Exploit (classified on the basis of their exploration–exploitation bias)

difference between their bias scores, $t[11] = 4.31, p < .005, d = 2.40$, and group Explore took more transitions to reach the critical decision than group Exploit, $t[11] = 2.82, p < .05, d = 1.61$.

7.4 | Supplementary analyses

Table 2 shows the Pearson's correlation coefficients between exploration loadings, exploitation loadings, the overall number of transition across the entire session and group size; remembering that while the exploration and exploitation loadings are derived from the nature of the transitions between SA, PF and PE, the number of transitions is coded independently of the nature of the transitions. Inspection of the table confirms that there is a similar pattern of correlations in the full set of 18 groups (values above the line of Xs) and the subset of 13 that came to the critical decision about evacuating local residents (values below the line of Xs). While exploration and exploitation loadings from the PCA (naturally) did not correlate with one another, both are correlated with the overall number of transitions, but neither correlated with group size. Thus, as the exploration and exploitation loadings increased, so the number of transitions (irrespective of their nature) increased; but the nature of the relationship between these loadings and the number of transitions per se did not differ. While there was some tendency for increases in group size to be related to reductions in both exploration and exploitation loadings, these relationships were not statistically significant. In summary, the results presented in Table 2 illustrate similarities between the full set of 18 groups and the subset of 13 that reached the critical decision, in terms of the relationships to the overall number of transitions within a session and group size. The results in Table 2 also show that the exploration and exploitation loadings are similarly related to the overall number of transitions and group size. Thus, the influence of exploration and exploitation on the number of transitions to the critical decision is unlikely to be a consequence of differences in these more general features of the groups.

8 | GENERAL DISCUSSION

Groups of individuals from different agencies come together to make life-determining decisions at major incidents. In the UK, they are

called SCGs. We investigated decision-making in such multi-agency groups during immersive simulations presented in Hydra suites (Alison et al., 2013; Crego, 1996). This approach—based on archival recordings of the same incidents being faced by different groups—enabled a systematic analysis of the process of group decision-making in this important context (see Wilkinson et al., 2019; see also, Waring et al., 2019; Waring et al., 2020). The recordings of the meetings were coded as sequences of decision-making activities that have been employed to characterize individual decision-making in the emergency services (i.e. situation assessment, SA; plan formulation, PF and plan execution, PE; see Burke et al., 2006; Cohen-Hatton et al., 2015; Cohen-Hatton & Honey, 2015; Lipshitz & Bar-Ilan, 1996; van den Heuvel *et al.*, 2012, van den Heuvel et al., 2014). This analysis allowed us to assess whether or not the groups followed a normative cyclical model of decision-making, which assumes that situation assessment is followed by plan formulation and then plan execution, and back to situation assessment. This could not have been achieved without the use of simulated environments, which enable replication, together with analysis of group decision processes in real-time. Our approach also enabled an assessment of between-group differences in styles of decision-making, which would not have been possible through studying a small number of groups or, in the limiting case, a single group (e.g. Waring et al., 2019; Waring et al., 2020; see also, Curnin et al., 2020).

Our overall analysis of the sequences of transitions revealed marked departures from the normative model. Over the course of the simulated incidents, situation assessment was more often followed by plan formulation than plan execution, but plan formulation was just as likely to be followed by situation assessment as it was to be followed by plan execution. Finally, plan execution was more likely to be followed by plan formulation than situation assessment. Further analysis of the sequences revealed that these departures from the normative standard involved systematic between-group differences in exploration and exploitation.

A principal-components analysis was conducted on the six pairwise transitions between the three decision-making activities (i.e. SA-PF, SA-PE, PF-PE, PF-SA, PE-SA, PE-PF). This analysis revealed that there were marked between-groups differences in the tendency to move between situation assessment and plan execution (i.e. SA-PF and PF-SA). At a descriptive level, these group differences could be aligned to differences in a process of decision inertia (Alison et al., 2015) or they could be characterized as reflecting differences

TABLE 2 Correlations between exploration, exploitation, number of transitions and group size

	Exploration	Exploitation	Number of transitions	Group size
Exploration	X	0	70**	-34
Exploitation	02	X	70**	-27
Number of transitions	71**	71**	X	-41
Group Size	-23	-25	-35	X

Note: Pearson's correlations above the Xs are from the 18 groups, while those below the Xs are from the 13 groups who reached the critical decision (to evacuate or not; * $p < .05$, ** $p < .01$; there were no correlations that were significant at the $p < .05$ level).

in the tendency to engage in exploration (cf. Bang & Frith, 2017). While decision inertia and exploration are conceptually distinct, they could prove difficult to tease apart. For example, repeated requests for (similar) information could be considered either decision inertia or exploration. However, our analysis also revealed group differences in the tendencies to move between plan execution to both situation assessment and plan formulation (i.e. PF–PE, PE–PF, PE–SA and SA–PE): A pattern of results that is indicative of exploiting information. Taken together, these between-group differences in the sequences of decision-making activities are clearly analogous to individual differences in information exploration and exploitation. We have already noted that there is evidence for individual differences in these processes (e.g. Badre *et al.*, 2012), and our results provide the first evidence that these differences can be seen at the level of group decision-making. Moreover, it mattered whether groups tended to explore or exploit: The tendency to explore (but not to exploit) was associated with a greater number of transitions (between decision-making activities) to reach a critical decision: whether or not to evacuate individuals from the location of the (simulated) major incident. To the extent that the increased number of transitions can be equated with the time at which critical decisions were made, then the consequences of groups tending to engage in exploration could be life determining.

Why do some groups explore and others exploit information? Our results are consistent with the general claim that the composition of groups might be an important determiner of effective group decision-making (Bang & Frith, 2017). The groups were opportunity samples of individuals from the various agencies, and this sampling approach mirrors how the composition of multi-agency groups who assemble to respond to major incidents in the UK might vary. If some individuals in the groups are explorers and others exploiters, then it is plausible to suppose that the relative numbers of these different individuals affect the behaviour of the groups. This analysis could be assessed by examining how such differences in group composition affect group decision-making. Alternatively, the decision-making style of specific individuals (e.g. the chair; see Wilkinson *et al.*, 2019; see also Waring *et al.*, 2020) might have a disproportionate influence on group decision-making processes: When the chair is an explorer (or exploiter) the group is more likely to exhibit the same bias. We have no evidence regarding the decision-making style of the chair that is independent of the decision-making processes in the group as a whole. However, individual variation in these decision-making biases could be assessed using a version of the “observe or bet” task prior to engaging in group decision-making (Tversky & Edwards, 1966). In any case, the results that we have presented take our understanding of group decision-making—at emergency incidents—from appeals to “groupthink” (see also Janis, 1972, 1982; Janis & Mann, 1977) to a form of analysis that is analytically tractable and testable (cf. Bang & Frith, 2017; see also, Alison *et al.*, 2015).

To conclude: The results from simulations provide one basis upon which to develop future policy, guidance and training for groups who have to make life-determining decisions under conditions of

uncertainty and time pressure. The analysis developed in the previous paragraphs highlights important issues around the selection of individuals to be charged with responding to major incidents. The suggestion that individuals could be selected on the basis on their individual decision-making styles (explorer or exploiters) is, at least in principle, a relatively simple one to implement (e.g. by using version of the “explore or bet” task). However, in situ monitoring and feedback of the utility of repeated requests for further information and additional situational assessment is another area that could be targeted in a variety of ways: by the presence of a critical friend or by training the chair to monitor the balance between exploration and exploitation (cf. Janis, 1972, 1982; Lovallo & Kahneman, 2003; Ministry of Defence, 2013; see also, Newell *et al.*, 2015). Finally, it should be acknowledged that our analysis is based on groups responding to simulated major incidents; albeit that the groups themselves are representative of groups who would be called to respond to real incidents. Future research will need to provide an experimental analysis of the origin of group differences in exploration and exploitation, for example, by manipulating the proportion of explorers and exploiters in different groups and examining the consequences for decision-making of this manipulation. This research too will need to be based on real groups responding to simulated major incidents.

ACKNOWLEDGEMENTS

B. Wilkinson was supported by a PhD studentship awarded by the School of Psychology. The development of the research was also supported by an Economic and Social Research Council Impact Acceleration Account fund, administered by Cardiff University (ES/M500422/1), awarded to R.C.H. and S.R.C-H. We thank the participants from Exercise Wales Gold 1 and 2, and our colleagues who contributed their time and expertise. B.W. was involved in all components of the research. B.W., S.R.C-H., and R.C.H. conceived the studies involving Exercise Wales Gold 1 and 2. B.W. and R.C.H. conducted the analysis of the results. B.W., S.R.C-H., and R.C.H. wrote the paper. The authors declare no conflicts of interest. The data reported in this paper will be archived on ORCA (Cardiff University) or made available on request for scholarly purposes.

REFERENCES

- Akaishi, R., Umeda, K., Nagase, A., & Sakai, K. (2014). Autonomous mechanisms of internal choice estimate underlies decision inertia. *Neuron*, *81*, 195–206.
- Alison, L., & Crego, J. (2008). *Policing major incidents: Leadership and major incident management*. Willan Publishing.
- Alison, L., Power, N., van den Heuvel, C., Humann, M., Palasinski, M., & Crego, J. (2015). Decision inertia: Deciding between least worst outcomes in emergency responses to disasters. *Journal of Occupational and Organizational Psychology*, *88*, 295–321. <https://doi.org/10.1111/joop.12108>
- Alison, L., van den Heuvel, C., Waring, S., Power, N., Long, A., O'Hara, T., & Crego, J. (2013). Immersive simulated learning environments (ISLEs) for researching critical incidents: A knowledge synthesis of the literature and experiences of studying high risk strategic and tactical decision making. *Journal of Cognitive Engineering and Decision Making*, *7*, 255–272.

- Badre D., Doll B. B., Long N. M., & Frank M. J. (2012). Rostrolateral prefrontal cortex and individual differences in uncertainty-driven exploration. *Neuron*, *73*, 595–607.
- Bang, D. & Frith, C. D. (2017). Making better decisions in groups. *Royal Society Open Science*, *4*, 170193. <https://doi.org/10.1098/rsos.170193>
- Burke, C. S., Stagl, K. C., Salas, E., Pierce, L., & Kendall, D. (2006). Understanding team adaptation: A conceptual analysis and model. *Journal of Applied Psychology*, *91*, 1189–1207. <https://doi.org/10.1037/0021-9010.91.6.1189>
- Cabinet Office. (2013). JESIP doctrine. Retrieved from <http://www.jesip.org.uk/uploads/resources/JESIP-Joint-Doctrine.pdf>
- Cohen J. D., McClure S. M., & Yu A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society London B, Biological Sciences*, *362*, 933–942.
- Cohen-Hatton, S. R., Butler, P. C., & Honey, R. C. (2015). An investigation of operational decision making in situ: Incident command in the UK fire and rescue service. *Human Factors*, *57*, 793–804. <https://doi.org/10.1177/0018720815578266>
- Cohen-Hatton, S. R. & Honey, R. C. (2015). Goal-oriented training affects decision-making processes in virtual and simulated fire and rescue environments. *Journal of Experimental Psychology: Applied*, *21*, 395–406.
- Comfort, L. K. (2007). Crisis management in hindsight: Cognition, communication, coordination and control. *Public Administration Review*, *67*, 189–197. <https://doi.org/10.1111/j.1540-6210.2007.00827.x>
- Crego, J. (1996). Critical incident management: Engendering experience through simulation. Unpublished PhD Thesis, University of Salford.
- Curnin, S., Brooks, B., & Owen, C. (2020). A case study of disaster decision-making in the presence of anomalies and absence of recognition. *Journal of Contingencies and Crisis Management*, *28*, 110–121. <https://doi.org/10.1111/1468-5973.12290>
- Daw N. D., O'Doherty J. P., Dayan P., Seymour B., & Dolan R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*.
- Dewey J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process*. Heath.
- Doya, K. (2008). Modulators of decision making. *Nature Neuroscience*, *11*, 410–416. <https://doi.org/10.1038/nn2077>
- Flin, R. (1996). *Sitting in the hot seat: Leaders and teams for critical incident management*. Wiley.
- Frank, M. J., Doll, B. B., Oas-Terpstra, J., & Moreno, F. (2009). Prefrontal and dopaminergic genes predict individual differences in exploration and exploitation. *Nature Neuroscience*, *12*, 1062–1068.
- Gigerenzer, G. (2007). *Gut feelings: The intelligence of the unconscious*. Viking.
- Groenendaal, J. & Helsloot, I. (2016). The application of naturalistic decision making (NDM) and other research: Lessons for frontline responders. *Journal of Management & Organization*, *22*, 173–185.
- Gureckis, T. M. & Goldstone, R. L. (2006). Thinking in groups. *Pragmatics & Cognition*, *14*, 293–311.
- House, A., Power, N., & Alison, L. (2014). A systematic review of the potential hurdles of interoperability to the emergency services in major incidents: Recommendations for solutions and alternatives. *Cognition, Technology & Work*, *16*, 319–335. <https://doi.org/10.1007/s10111-013-0259-6>
- Hutchins, E. (1995a). How a cockpit remembers its speed. *Cognitive Science*, *19*, 265–288.
- Hutchins, E. (1995b). *Cognition in the wild*. MIT Press. Retrieved from <https://www.hydrfoundation.org>
- Janis, I. L. (1972). *Victims of groupthink*. Houghton Mifflin.
- Janis, I. L. (1982). *Groupthink: Psychological studies of policy decisions and fiascoes*, 2nd ed. Houghton Mifflin.
- Janis, I. L. & Mann, L. (1977). Emergency decision making: A theoretical analysis of responses to disaster warnings. *Journal of Human Stress*, *3*, 35–48. <https://doi.org/10.1080/0097840X.1977.9936085>
- Klein, G. (1993). A recognition-primed decision (RPD) model of rapid decision making. In G. Klein, J. Orasanu, R. Calderwood, & C. Zsombok (Eds.), *Decision making in action* (pp. 138–147). Ablex.
- Klein, G. (2003). *Intuition at work*. Doubleday.
- Klein, G. A. (2008). Naturalistic decision making. *Human Factors*, *50*, 456–460. <https://doi.org/10.1518/001872008X288385>
- Klein, G., Ross, K. G., Moon, B. M., Klein, D. E., Hoffman, R. R., & Hollnagel, E. (2003). Macrocognition. *IEEE Intelligent Systems*, *18*, 81–85. <https://doi.org/10.1109/MIS.2003.1200735>
- Klein G. (1998). *Sources of Power: How people make decisions*. MIT Press.
- Klein G. A., Calderwood R., & MacGregor D. (1989). Critical decision method for eliciting knowledge. *Systems, Man and Cybernetics, IEEE Transactions on*, *19*, 462–472.
- Lipshitz, R. & Bar-Ilan, O. (1996). How problems are solved: Reconsidering the phase theorem. *Organizational Behavior and Human Decision Processes*, *65*, 48–60.
- Lovullo, D. & Kahneman, D. (2003). Delusions of success. *Harvard Business Review*, *81*, 57–63.
- Ministry of Defence. (2013). *Red teaming guide*, 2nd ed. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/142533/20130301_red_teaming_ed2.pdf
- National Decision Model. Retrieved from <https://www.app.college.police.uk/app-content/national-decision-model/the-national-decision-model/#the-model>
- Newell, B. R., Lagnado, D. A., & Shanks, D. R. (2015). *Straight choices: The psychology of decision making*, 2nd ed. Psychology Press.
- O'Connor, B. P. (1999). Simple and flexible SAS and SPSS programs for analyzing lag-sequential categorical data. *Behavior Research Methods, Instruments, and Computers*, *31*, 718–726.
- Pollock, K. (2013). *Review of persistent lessons identified relating to interoperability from emergencies and major incidents since 1986*. Occasional Paper 6, Emergency Planning College, Easingwold.
- Polzer, J. T., Milton, L. P., & Swann, W. B. (2002). Capitalizing on diversity: Interpersonal congruence in small work groups. *Administrative Science Quarterly*, *47*, 296–324.
- Power, N. & Alison, L. (2017). Offence or defence?: Approach and avoid goals in the multi-agency emergency response to a simulated terrorism attack. *Journal of Occupational and Organizational Psychology*, *90*, 51–76.
- Roberge, M. E. & van Dick, R. (2010). Recognizing the benefits of diversity: When and how does diversity increase group performance? *Human Resource Management Review*, *20*, 295–308.
- Sackett, P. (1979). The lag sequential analysis of contingency and cyclicity in behavioural interactional research. In J. D. Osofsky (Ed.), *Handbook of infant development* (pp. 623–649). Wiley and Sons.
- Salas, E., DiazGranados, D., & Rosen, M. A. (2010). Expertise-based intuition and decision making in organizations. *Journal of Management*, *36*, 941–973. <https://doi.org/10.1177/0149206309350084>
- Tversky, A. & Edwards, W. (1966). Information versus reward in binary choices. *Journal of Experimental Psychology*, *71*, 680–683. <https://doi.org/10.1037/h0023123>
- Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131. <https://doi.org/10.1126/science.185.4157.1124>
- van den Heuvel, C., Alison, L., & Power, N. (2014). Coping with uncertainty: Police strategies for resilient decision-making and action implementation. *Cognition, Technology & Work*, *16*, 25–45. <https://doi.org/10.1007/s10111-012-0241-8>
- Van den Heuvel C., Alison L., & Crego J. (2012). How uncertainty and accountability can derail strategic 'save life' decision in counterterrorism simulations: a descriptive model of choice deferral and omission bias. *Journal of Behavioral Decision Making*, *25*, 165–187.
- Waring, S., Alison, L., Shortland, N., & Human, M. (2019). The role of information sharing on decision delay during multiteam disaster

response. *Cognition, Technology & Work*, 22(2), 263–279. <https://doi.org/10.1007/s10111-019-00570-7>

Waring, S., Moran, J.-L., & Page, R. (2020). Decision-making in multi-agency multiteam systems operating in extreme environments. *Journal of Occupational and Organizational Psychology*, 93, 629–653. <https://doi.org/10.1111/joop.12309>

Wilkinson, B., Cohen-Hatton, S. R., & Honey, R. C. (2019). Decision making in multi-agency groups at simulated major incident emergencies: In situ analysis of adherence to UK doctrine. *Journal of Contingencies and Crisis Management*, 27, 306–316. <https://doi.org/10.1111/1468-5973.12260>

How to cite this article: Wilkinson B, Cohen-Hatton SR, Honey RC. Variation in exploration and exploitation in group decision-making: Evidence from immersive simulations of major incident emergencies. *J Contingencies and Crisis Management*. 2021;00:1–10. <https://doi.org/10.1111/1468-5973.12355>